

FIGURE 1

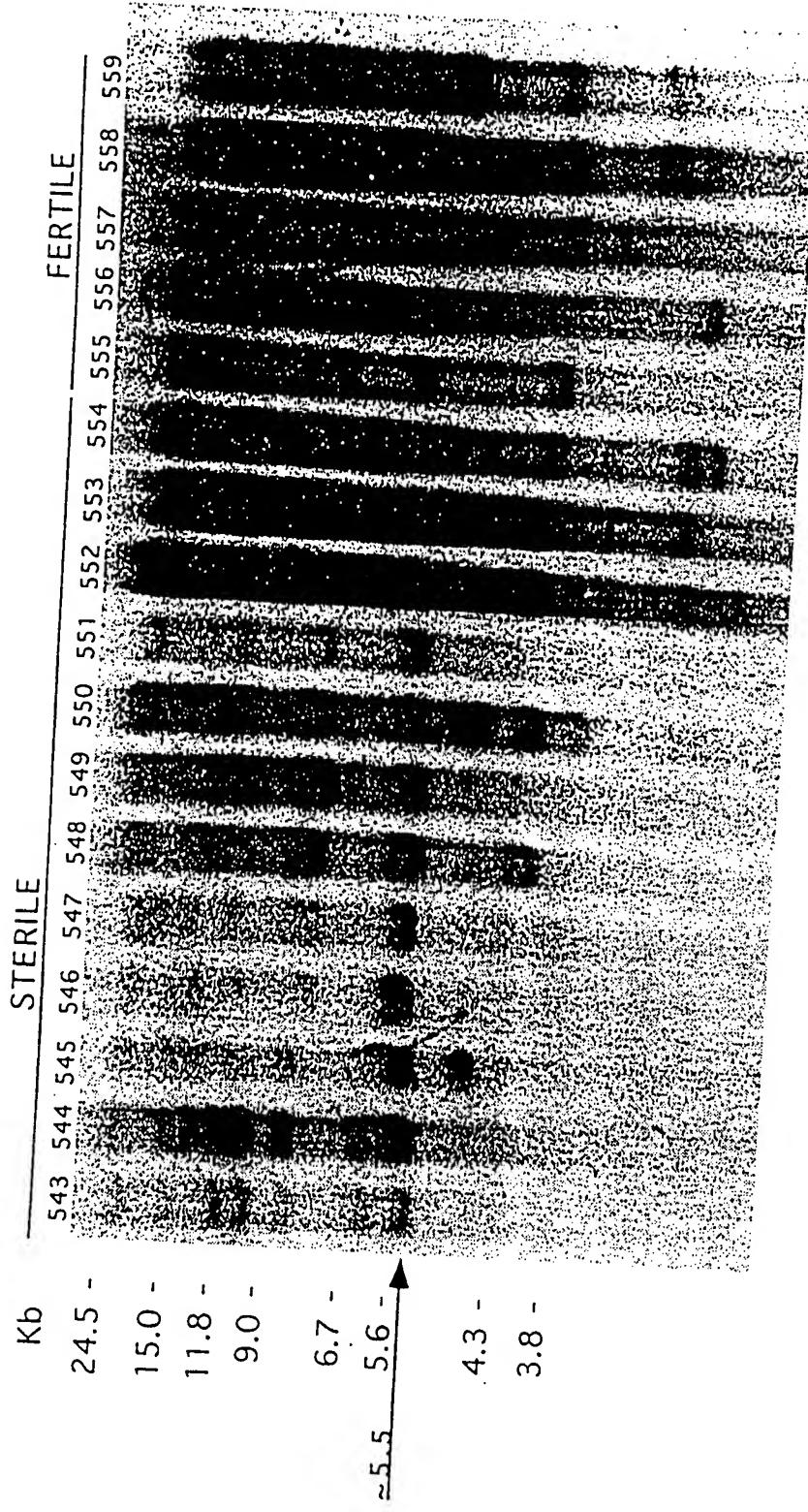


FIGURE 2

00024652 2000

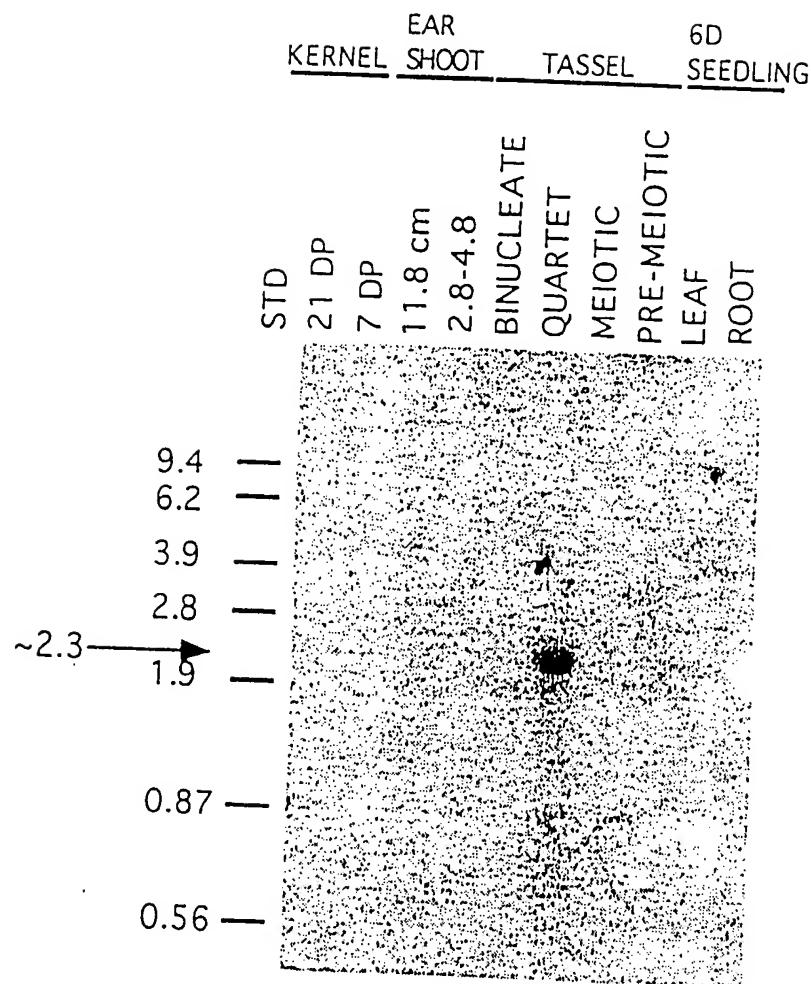


FIGURE 3

Figure 4

ECORI

1 GAATTGGCACGAGGGAAAGCTCACCTCACGCCGGCACGCCATGCCATTCTTCCACTA
 1 CTTAACCGTGCTCCCTCGAGTGGAGTGCAGGCGCTGCGGTAGCGGTAAAGAAGGGTGT 60

a E F G T R E A H L T P A T P P S P F F P L -
 61 GCAGGGCCTCACAAAGTACATCGCGCTCCTCTGGTTGTCCTCTCATGGATCCTGGTCCAG
 61 CGTCCCGAGTGTTCATGTAGCGCGAGGAAGACCAACAGGAGAGTACCTAGGACCAAGGTC 120

a A G P H K Y I A L L L V V L S W I L V Q -
 121 AGGTGGAGCCTGAGGAAGCAGAAAGGCCGAGATCATGCCAGTCATCGGCCAACGGTG
 121 TCCACCTCGGACTCCTCGTCTTCCGGCTCTAGTACCGGTAGTAGCCCGTTGCCAC 180

a R W S L R K Q K G P R S W P V I G A T V -
 181 GAGCAGCTGAGGAACTACCACCGGATGCACGACTGGCTTGTGGGTACCTGTACGGCAC
 181 CTCGTCGACTCCTTGATGGTGGCCTACGTGCTGACCGAACAGCCATGGACAGTGCCGTG 240

a E Q L R N Y H R M H D W L V G Y L S R H -
 241 AGGACAGTGACCGTCGACATGCCGTTCACTTCCTACACCTACATCGCTGACCCGGTGAAT
 241 TCCTGTCACTGGCAGCTGTACGGCAAGTGAAGGATGTGGATGTAGCGACTGGCCACTTA 300

a R T V T V D M P F T S Y T Y I A D P V N -
 301 GTCGAGCATGCTCAAGACTAACTTCACCAATTACCCAAGGGATCGTGTACAGATCC
 301 CAGCTCGTACAGGAGTTCTGATTGAAGTGGTTAATGGGGTTCCCTAGCACATGTCTAGG 360

a V E H V L K T N F T N Y P K G I V Y R S -
 361 TACATGGACGTGCTCCTCGGTGACGGCATCTCAACGCCGACGGCGAGCTGTGGAGGAAG
 361 ATGTACCTGCACGAGGAGCCACTGCCGTAGAAGTTGGCTGCCGCTCGACACCTCCCTC 420

a Y M D V L L G D G I F N A D G E L W R K -
 421 CAGAGGAAGACGGCGAGTTCGAGTTCGCCTCCAAGAACCTGAGGGATTCAGCGCCATT
 421 GTCTCCTCTGCCGCTCAAAGCTCAAGCGGAGGTTCTGGACTCCCTAAAGTCGCGGTAA 480

a Q R K T A S F E F A S K N L R D F S A I -

Figure 4B

GTGTTCAGAGAGTACTCCCTGAAGCTGTCGGGTACTGAGCCAGGCATCCAAGGCAGGC
 481 CACAAGTCTCATGAGGGACTTCGACAGCCCATATGACTCGGTCCTAGGTTCCGTCCG 540
 a V F R E Y S L K L S G I L S Q A S K A G -
 AAAGTTGTGGACATGCAGGAACCTTACATGAGGATGACGCTGGACTCCATCTGCAAGGTT
 541 TTTCAACACCTGTACGTCCCTGAAATGTACTCCTACTGCGACCTGAGGTAGACGTTCCAA 600
 a K V V D M Q E L Y M R M T L D S I C K V -
 GGGTTCGGGGTGAGATCGGCACGCTGTCGCCAGATCTCCCCGAGAACAGCTTCGCGCAG
 601 CCCAAGCCCCAGCTCTAGCGTGCACAGCGGTCTAGAGGGGCTCTGTCGAAGCGCGTC 660
 a G F G V E I G T L S P D L P E N S F A Q -
 GCGTTCGATGCCGCCAACATCATCATCACGCTGCGGTTATCGACCCGCTGTGGCGCATC
 661 CGCAAGCTACGGCGGTTGTAGTAGTAGTGCACGCCAAGTAGCTGGCGACACCGCGTAG 720
 a A F D A A N I I I T L R F I D P L W R I -
 AAGAGGTTCTTCCACGTGGTCAGAGGCCCTCCTAGCGCAGAGCATCAAGCTCGTGGAC
 721 TTCTCCAAGAAGGTGCAGCCCAGTCTCGGGAGGATCGCGTCTCGTAGTTCGAGCACCTG 780
 a K R F F H V G S E A L L A Q S I K L V D -
 GAGTTCACCTACAGCGTGTCCGGAGGAAGGCCAGATCGTCGAGGTCCGGCCAGC
 781 CTCAAGTGGATGTCGCACTAGGCGGCCCTCCGGCTCTAGCAGCTCCAGGCCGGTCG 840
 a E F T Y S V I R R R K A E I V E V R A S -
 GGCAAACAGGAGAAGATGAAGCACGACATCCTGTCACGGTTATCGAGCTGGCGAGGCC
 841 CCGTTTGTCCCTTCTACTTCGTGCTGTAGGACAGTGCACAGTAGCTCGACCCGCTCCGG 900
 a G K Q E K M K H D I L S R F I E L G E A -
 GGCGACGACGGCGGCGGCTTCGGGACGATAAGAGCCTCCGGACGTGGTGCTCAACTTC
 901 CCGCTGCTGCCGCCGAAGCCCTGCTATTCTGGAGGCCCTGCACACGAGTTGAAG 960
 a G D D G G G F G D D K S L R D V V L N F -
 GTGATCGCCGGCGGGACACGACGGCGACGACGCTGTCGTGGTTACCGCACATGCCATG
 961 CACTAGCGGCCGCCCTGTGCTGCCGCTGCGACAGCACCAAGTGCCTGACCGGTAC 1020

Figure 4C

1021 TCCCACCCGGACGTGGCCGAGAAGCTGCGCCGAGCTGTGCGCGTTCGAGGCGGAGCGC
 1080 AGGGTGGGCTGCACCGGCTTCGACGCGCGCTCGACACGCGAAGCTCCGCCTCGCG 1080
 a S H P D V A E K L R R E L C A F E A E R -
 1081 GCGCGCGAGGAGGGCGTCACGCTCGTCTGCGGCGCGCTGACGCCGACGACAAGGCG
 1140 CGCGCGCTCCTCCCGCAGTGCAGCAGACGCCGCGACTGCGGCTGCTGTTCCGC 1140
 a A R E E G V T L V L C G G A D A D D D K A -
 1141 TTGCGCCCGCGTGGCGCAGTCGCGGGCTCCTCACCTACGACAGCCTCGCAAGCTG
 1200 AAGCGGCGGGCGCACCGCGTCAAGCGCCGGAGGAGTGGATGCTGCGAGGCCGTCGAC 1200
 a F A A R V A Q F A G L L T Y D S L G K L -
 1201 GTCTACCTCACGCCGTCACCGAGACGCTCCGCTGTACCCGCCGTCCTCAGGAC
 1260 CAGATGGAGGTGGACGAGTGGCTCTGCGAGGCGGACATGGGGCGGAGGGAGTCCTG 1260
 a V Y L H A C V T E T L R L Y P A V P Q D -
 1261 CCCAAGGGGATCCTGGAGGACGACGTGCTGCCGGACGGACGAAGGTGAGGGCCGGCGG
 1320 GGGTTCCCTAGGACCTCCTGCTGCACGACGCCGTCCTGCTTCACTCCCGCCGCC 1320
 a P K G I L E D D V L P D G T K V R A G G -
 1321 ATGGTGACGTACGTGCCCTACTCGATGGGGCGGATGGAGTACAACGGGGCCCCGACGCG
 1380 TACCACTGCATGCACGGGATGAGCTACCCGCCTACCTCATGTTGACCCGGGGCTGCGC 1380
 a M V T Y V P Y S M G R M E Y N W G P D A -
 1381 GCGAGCTCCGGCGGAGCGGTGGATCAACGAGGATGGCGCTCCGCAACGCGTCGCCG
 1440 CGCTCGAAGGCCGCGCTGCCACCTAGTTGCTCCTACCGCGCAAGGCGTTGCGCAGCGC 1440
 a A S F R P E R W I N E D G A F R N A S P -
 1441 TTCAAGTTCACGGCGTTCAAGGCGGGCCGAGGATCTGCCTGGCAAGGACTCGCGTAC
 1500 AAGTTCAAGTGGCGCAAGGTCCGCCCGCTCCTAGACGGACCCGTTCTGAGCCGCATG 1500
 a F K F T A F Q A G P R I C L G K D S A Y -
 1501 CTGCAGATGAAGATGGCGCTGGCCATCTCTTCCGTTCTACAGCTCCGGCTGCTGGAG
 1560 GACGTCTACTTCTACCGCGACCGGTAGGAGAAGGCGAAGATGTCGAAGGCCGACGACCTC 1560
 a L Q M K M A L A I L F R F Y S F R L L E -

Figure 4D

1561 GGGCACCCGGTGCAGTACCGCATGATGACCATCCTCTCCATGGCGCACGGCCTCAAGGTC
 CCCGTGGGCCACGTCTGGCGTACTACTGGTAGGAGAGTACCGCGTCCGGAGTTCCAG 1620
 G H P V Q Y R M M T I L S M A H G L K V -
 CGCGTCTCTAGGGCCGCTCTGATGTCATGGCGATTGGATATGGATATCGTCCCGCTTAAT
 1621 GCGCAGAGATCCCGCAGACTACAGTACCGCTAAACCTATACCTATAGCAGGGCGAATT 1680
 R V S R A V * C H G D L D M D I V P L N -
 CCACGACAAATAACGCTCGTGTACAAATTGCATGCATGCATGTAAGGGAAAGCGATGG
 1681 GGTGCTGTTATTGCGAGCACAATGTTAACGTACGTACGTACATTCCCTTCGCTACC 1740
 P R Q I T L V L Q I C M H A C K G K R W -
 GTTCAATTGGTGGCTTGGCTTAAGCCTTAAAAACTCCGTCGGGTCTTGCAGACCCACACA
 1741 CAAAGTAACCACCGAACCGAATTCCGAATTGGAGGCAAGCCAGAACGCTTGGTGGTGT 1800
 V S L V A W L K P *
 TCACTAGTGTGGCTACTCTACTCCTCAGTGGAAAGTGTAGTGACAGCATAAAGTTCATC
 801 AGTGATCACAAAACATGAGATGAGGAGTCACCTCACATCACTGTCGTATGTTCAAGTAG 1860

XbaI

|

1861 ATATATATTATCCTCTTCTTAAAAA~~AA~~ACTCGAG
-----+-----+-----+-----+-----+-----+----- 1906
TATATATAATAGGAGAAAGAATTTTTTTTTTTTGAGCTC

Figure 5

1 GAATTCCAAG CGAGGCCCTT GTAGCAGAGA GTGTTGCTGA TGCAGTCGGC
51 GGAAATGAGT GCGTGCTGAG AGCAACGCTG AGGGGTTCCA GGGATGGCAA
101 TGGCTATGGC AATCGGCTAG AGGTGGAGGA CAAGGTGGTG AGGATTGGGA
151 GGGCAACCTA TGGCAAGTTG GTGAAGAGGC ACGCAATGAG AGATCTATTG
201 AGACTTACAC TGGATGCCGC CAACAAATTG AACCTTTAGA TTTTGATACT
251 GTCACTCCTA CTTTATTCCCT TGGTTGGCA ACTTCCAATA GGCTCATGTT
301 AATCAATGAT TAGTGATTAT TCAGCAAATA TTCTTGTTG TTTGACATTT
351 ATAATATGTG GGGTGAGACG GATTAATAT CATCCATGAG AGCTTTATCT
401 TCATGCTCTC TTGATTTGG TTTCAGATCA TTCTTCAGT GTTCACAAGA
451 ATTTTCTCAG TTTGGTCCAT GTAATTTTG AAGTGAGGTT CCTTAAATTT
501 CATTATGCTT CCTTTCTTT CTAGACTAGC AACTGCATGA CTTTCACCT
551 TGGGTTCAC AATTGACTCA CAAGAAAACA AATTCACTTT TGGGTTCAC
601 AATTCCCTTT CAGGATGTAC TTTTCACTTG AACTGTCATG TATAGGAACA
651 AGGAATGGCT CAGTTTTAA GGAACAATGT ACAGATTTCA TTTCAGAACT
701 CTTTCTGGTT GGTTGAGTTT CAGACTTTT GTACCAAGCT GATGGATCAC
751 AATACTTGTT TCCAAAGTCT GATAACAGAA ACTGGCAACT CCTAATTGAT
801 AATAAAAAGA ATAAAATACA GTATCAGATA TCTCATTTC TTGGTTGGCA
851 GATCACAAAA AGGAACACAA AGGCTAAGCC TCCTACTTGT TCGGGAGTTA
901 GGTCAGGGAC ACCATATGAA TGAAAGAAAT CTTAATTGG GGTACACACCA
951 AGATTGTCTC TCTCGAGGTT GGGGGTCCC TAAGGTTGGT AGTAGCAATA
1001 CCCAATATAT CACCTAACAA ACCCAATCCA TGCTACATAC ATACATAGCA
1051 TCCATCACTT GTAGACTGGA CCCTTCATCA AGAGCACCCT GGAGGAAGCT
1101 CACATCACGC CGGCGACGCC ATCGCCATTC TTCCCCTAG CAGGGCCTCA
1151 CAAGTACATC GCGCTCCTCC TGGTTGTCCT CTCATGGATC CTGGTCCAGA
1201 GGTGGAGCCT GAGGAAGCAG AAAGGCCGA GATCATGGCC AGTCATCGGT
1251 GCAACGGTGG AGCAGCTGAG GAACTACCAC CGGATGCACCG ACTGGCTTGT
1301 CGGGTACCTG TCACGGCACA GGACAGTGAC CGTCGACATG CCGTTCACTT
1351 CCTACACCTA CATCGCTGAC CCGGTGAATG TCGAGCATGT CCTCAAGACT

10021652-121404

Figure 5B

1401 AACTTCACCA ATTACCCCAA GGTAAATGAC CTGAACTCAC TGATGTTAG
1451 TCTTCGGAAA TCAGAGCTGA AAGCTGAATC GAATGTGCCT GAACACCGTG
1501 TAGGGAATCG TGTACAGATC CTACATGGAC GTGCTCCTCG GTGACGGCAT
1551 CTTCAACGCC GACGGCGAGC TGTGGAGGAA GCAGAGGAAG ACGCGAGTT
1601 TCGAGTTCGC CTCCAAGAAC CTGAGGGATT TCAGCGCCAT TGTGTTCAGA
1651 GAGTACTCCC TGAAGCTGTC GGGTATACTG AGCCAGGCAT CCAAGGCAGG
1701 CAAAGTTGTG GACATGCAGG TGAGATCACT GCTCCCTTGC CATTGCCAAC
1751 ATGAGCATT CAACCTGAGA CACGAGAGCT ACCTTGCCGA TTCAGGAACT
1801 TTACATGAGG ATGACGCTGG ACTCCATCTG CAAGGTTGGG TTCGGGGTCG
1851 AGATCGGCAC GCTGTCGCCG GATCTCCCCG AGAACAGCTT CGCGCAGGCG
1901 TTTCGATGCCG CCAACATCAT CGTCACGCTG CGGTTCATCG ACCCGCTGTG
1951 GCGCATCAAG AGGTTCTTCC ACGTCGGGTC AGAGGCCCTC CTAGCGCAGA
2001 GCATCAAGCT CGTGGACGAG TTCACCTACA GCGTGATCCG CCGGAGGAAG
2051 GCCGAGATCG TCGAGGCCCG GGCCAGCGGC AAACAGGAGA AGGTACGTGC
2101 ACATGACTGT TTTCGATTCTT CAGTTCATCG TCTTGGCCCG GATGGACCTG
2151 ATCCTGATTG ATTATATATC CGTGTGACTT GTGAGGACAA ATTAAAATGG
2201 GCAGATGAAG CACGACATCC TGTCACGGTT CATCGAGCTA GGCGAGGCCG
2251 GCGACGACGG CGGCGGCTTC GGGGACGACA AGAGCCTCCG GGACGTGGTG
2301 CTCAACTTCG TGATGCCGG GCGGGACACG ACGGCGACGA CGCTGTCGTG
2351 GTTCACGCAC ATGGCCATGT CCCACCCGGA CGTGGCCGAG AAGCTGCGCC
2401 GCGAGCTGTG CGCGTTCGAG GCGGAGCGCG CGCGCGAGGA GGGCGTCGCG
2451 CTCGTGCCCT GCGGCCGCG TGACGCCGAC GACAAGGCCT TCGCCGCCCG
2501 CGTGGCGCAG TTCGCGGGCC TCCTCACCTA CGACAGCCTC GGCAAGCTGG
2551 TCTACCTCCA CGCCTGCGTC ACCGAGACGC TCCGCCTGTA CCCCAGCGTC
2601 CCTCAGGTGA GCGCGCCCGA CACGCGACCT CCGGTCCAGA GCACAGCATG
2651 CAGTGAGTGG ACCTGAATGC AATGCACATG CACTTGCGCG CGCGCAGGAC
2701 CCCAAGGGGA TCCTGGAGGA CGACGTGCTG CCGGACGGGA CGAAGGTGAG
2751 GGCGGGCGGG ATGGTGACGT ACGTGCCCTA CTCGATGGGG CGGATGGAGT

10024657-242404

Figure 5C

2801 ACAACTGGGG CCCCCACGCG GCGAGCTTCC GGCCGGAGCG GTGGATCAAC
2851 GAGGATGGCG CGTTCCGCAA CGCGTCGCCG TTCAAGTTCA CGGCCTTCCA
2901 GGCGGGGCCG AGGATCTGCC TGGGCAAGGA CTCGGCGTAC CTGCAGATGA
2951 AGATGGCGCT GGCCATCCTC TTGCGCTTCT ACAGCTTCCG GCTGCTGGAG
3001 GGGCACCCGG TGCAGTACCG CATGATGACC ATCCTCTCCA TGGCGCACGG
3051 CCTCAAGGTC CGCGTCTCTA GGGCCGTCTG ATGTCATGGC GATTGGGAT
3101 ATCATCCCAC TTAATCCTTA AAAATTGCA TGCATGCATG TAAGGGAAAG
3151 CGATGGGTTT CATTGGTGGC TTGGCTTAAG CCTTAAAAAC TCCGTCGGGT
3201 CTTGCGAACCC ACCACATCAC TAGTGTGGT TACTCTACTC CTCAGTGGAA
3251 GTGTAGTGAC AGCATAACAAG TTCATCATAT ATATTATCCT CTTTCTTCGC
3301 CGGATGCTTC CCGGGACCTT TTGGAGACCA TTACTGACAG GCGTGTGAAA
3351 AAAAGGCTTC TTCTGCGGCG AAGTTTGGG TTCAGAGTCT TGGCGTCTTT
3401 GCAGCAGAAA AAAGGTTGG AAGGATCTGA ACCCTGAACC GAAAATGGCT
3451 TCGGAAATAT GCTCGCATCG GGGCGGGGCC GTCACTCGGG ATGACGACAA
3501 GCCCACAAAGC AGTGAGAGCG AAGCGATCTT TGGAGTTGG AGACACTCTC
3551 GGACCCCTCG GCGCTCCGCG AGCTCATCTT CGCCTCCTCT GTCGTGTCCG
3601 TGGCGGCACC GCGCCCGCCC GCCTCGTGT CGACCAAATC CCGCGCCCCG
3651 ACCGGTTCGT GTACAACACC CTCATCCGCG GCGCCGCGCG CAGTGACACG
3701 CCCCCGGACG CCGTATACAT CTATAATCA TGGTATTGTA CTTTATTTTC
3751 AAACGGCCTT AACACAAACCA TATTTTATG GTAAACACGT TCAAAATTGA
3801 CACAAATTAA AACACAGGCAC AAACCGTAGC TAAACATAAG AGAATGAGAG
3851 ACAACCCAAA GGTTAGAGAT GAAATAAGCT GAGTAAACGA CGAATTC

Figure 6

1051 TCCATCACTTGTAGACTGGACCCTTCATCAAGAGCACCATGGAGGAAGCT 1100
1 GAATTGGCACGGAGGAAGCT 21

1101 CACATCACGCCGGCGACGCCATGCCATTCTCCACTAGCAGGGCCTCA 1150
1122 CACCTCACGCCGGCGACGCCATGCCATTCTCCACTAGCAGGGCCTCA 71

1151 CAAGTACATCGCGCTCCTCTGGTTGTCCTCATGGATCCTGGTCCAGA 1200
1172 CAAGTACATCGCGCTCCTCTGGTTGTCCTCATGGATCCTGGTCCAGA 121

1201 GGTGGAGCCTGAGGAAGCAGAAAGGCCGAGATCATGCCAGTCATCGGT 1250
122 GGTGGAGCCTGAGGAAGCAGAAAGGCCGAGATCATGCCAGTCATCGGC 171

1251 GCAACGGTGGAGCAGCTGAGGAACTACCACCGGATGCACGACTGGCTTGT 1300
1272 GCAACGGTGGAGCAGCTGAGGAACTACCACCGGATGCACGACTGGCTTGT 221

1301 CGGGTACCTGTACGGCACAGGACAGTGACCGTCGACATGCCGTTCACTT 1350
1322 CGGGTACCTGTACGGCACAGGACAGTGACCGTCGACATGCCGTTCACTT 271

1351 CCTACACCTACATCGCTGACCCGGTGAATGTCGAGCATGTCCCTCAAGACT 1400
1372 CCTACACCTACATCGCTGACCCGGTGAATGTCGAGCATGTCCCTCAAGACT 321

1401 AACTCACCAATTACCCCAAGGTAAATGACCTGAACTCACTGATGTTCAG 1450
1422 AACTCACCAATTACCCCA..... 340

1501 TAGGGAATCGTGTACAGATCCTACATGGACGTGCTCCTCGGTGACGGCAT 1550
1522 AGGGATCGTGTACAGATCCTACATGGACGTGCTCCTCGGTGACGGCAT 389

1551 CTTCAACGCCGACGGCGAGCTGTGGAGGAAGCAGAGGAAGACGGCGAGTT 1600
1570 CTTCAACGCCGACGGCGAGCTGTGGAGGAAGCAGAGGAAGACGGCGAGTT 439

1601 TCGAGTTGCCCTCCAAGAACCTGAGGGATTCAGGCCATTGTGTTCAGA 1650
1640 TCGAGTTGCCCTCCAAGAACCTGAGGGATTCAGGCCATTGTGTTCAGA 489

1651 GAGTACTCCCTGAAGCTGTCGGGTATACTGAGCCAGGCATCCAAGGCAGG 1700
1690 GAGTACTCCCTGAAGCTGTCGGGTATACTGAGCCAGGCATCCAAGGCAGG 539

1701 CAAAGTTGTGGACATGCAGGTGAGATCACTGCTCCCTGCCATTGCCAAC 1750
1740 CAAAGTTGTGGACATG..... 555

Figure 6B

1751 ATGAGCATTCAACCTGAGACACGAGAGCTACCTTGCCTGCGATTCAAGGAAC 1800
556 CAGGAAC 563
|||
1801 TTACATGAGGATGACGCTGGACTCCATCTGCAAGGTTGGGTCGGGTG 1850
|||
564 TTACATGAGGATGACGCTGGACTCCATCTGCAAGGTTGGGTCGGGTG 613
|||
1851 AGATCGGCACGCTGTCGCCGGATCTCCCCGAGAACAGCTTCGCGCAGGCG 1900
|||
614 AGATCGGCACGCTGTCGCCAGATCTCCCCGAGAACAGCTTCGCGCAGGCG 663
|||
1901 TTGATGCCGCCAACATCATCGTCACGCTGCGGTTCATCGACCCGCTGTG 1950
|||
664 TTGATGCCGCCAACATCATCACCGCTGCGGTTCATCGACCCGCTGTG 713
|||
1951 GCGCATCAAGAGGTTCTTCCACGTCGGGTAGAGGCCCTCTAGCGCAGA 2000
|||
714 GCGCATCAAGAGGTTCTTCCACGTCGGGTAGAGGCCCTCTAGCGCAGA 763
|||
2001 GCATCAAGCTCGTGGACGAGTTCACCTACAGCGTATCCGCCGGAGGAAG 2050
|||
764 GCATCAAGCTCGTGGACGAGTTCACCTACAGCGTATCCGCCGGAGGAAG 813
|||
2051 GCCGAGATCGTCGAGGCCCGGGCCAGCGGAAACAGGAGAAGGTACGTGC 2100
|||
814 GCCGAGATCGTCGAGGTCCGGGCCAGCGGAAACAGGAGA..... 853
|||
2201 GCAGATGAAGCACGACATCCTGTCACGGTTCATCGAGCTAGGGAGGCCG 2250
|||
854 ..AGATGAAGCACGACATCCTGTCACGGTTCATCGAGCTGGCGAGGCCG 901
|||
2251 GCGACGACGGCGGGCTTCGGGACGACAAGAGCCTCCGGACGTGGTG 2300
|||
902 GCGACGACGGCGGGCTTCGGGACGATAAGAGCCTCCGGACGTGGTG 951
|||
2301 CTCAACTCGTGATCGCCGGCGGGACACGACGGGACGACGCTGTG 2350
|||
952 CTCAACTCGTGATCGCCGGCGGGACACGACGGGACGACGCTGTG 1001
|||
2351 GTTCACGCACATGGCCATGTCCCACCCGGACGTGGCGAGAACGCTGCC 2400
|||
1002 GTTCACGCACATGGCCATGTCCCACCCGGACGTGGCGAGAACGCTGCC 1051
|||
2401 GCGAGCTGTGCGCGTTCGAGGGGGAGCGCGCGCGAGGAGGGCGTCGCG 2450
|||
1052 GCGAGCTGTGCGCGTTCGAGGGGGAGCGCGCGCGAGGAGGGCGTCACG 1101
|||
2451 CTCGTGCCCTGCGCGGGCGCTGACGCCACGACAAGGCCTGCCGCCG 2500
|||
1102 CTCGTGCTCTGCGCGCGCTGACGCCACGACAAGGCCTGCCGCCG 1151

Figure 6C

2501 CGTGGCGCAGTCGGGGCTCCTCACCTACGACAGCCTCGGCAAGCTGG 2550
1152 CGTGGCGCAGTCGGGGCTCCTCACCTACGACAGCCTCGGCAAGCTGG 1201
2551 TCTACCTCCACGCCCTGCGTACCGAGACGCTCCGCCGTACCCGCCGTC 2600
1202 TCTACCTCCACGCCCTGCGTACCGAGACGCTCCGCCGTACCCGCCGTC 1251
2601 CCTCAGGTGAGCGCGCCGACACGCGACCTCCGGTCCAGAGCACAGCATG 2650
1252 CCT..... 1254
2651 CAGTGAGTGGACCTGAATGCAATGCACATGCACTTGCAGCGCCAGGAC 2700
1255 CAGGAC 1260
2701 CCCAAGGGGATCTGGAGGACGACGTGCTGCCGACGGACGAAGGTGAG 2750
1261 CCCAAGGGGATCTGGAGGACGACGTGCTGCCGACGGACGAAGGTGAG 1310
2751 GGCCGGCGGGATGGTGACGTACGTGCCCTACTCGATGGGGCGGATGGAGT 2800
1311 GGCCGGCGGGATGGTGACGTACGTGCCCTACTCGATGGGGCGGATGGAGT 1360
2801 ACAACTGGGGCCCCGACCGCGAGCTCCGGCCGGAGCGGTGGATCAAC 2850
1361 ACAACTGGGGCCCCGACCGCGAGCTCCGGCCGGAGCGGTGGATCAAC 1410
2851 GAGGATGGCGCGTCCGCAACCGTCGCCGTTCAAGTTACGGCGTTCCA 2900
1411 GAGGATGGCGCGTCCGCAACCGTCGCCGTTCAAGTTACGGCGTTCCA 1460
2901 GGCGGGGCCGAGGATCTGCCCTGGCAAGGACTCGGCGTACCTGCAGATGA 2950
1461 GGCGGGGCCGAGGATCTGCCCTGGCAAGGACTCGGCGTACCTGCAGATGA 1510
2951 AGATGGCGCTGGCCATCCTCTTGCCTCTACAGCTTCCGGCTGCTGGAG 3000
1511 AGATGGCGCTGGCCATCCTCTTCCGCTTACAGCTTCCGGCTGCTGGAG 1560
3001 GGGCACCCGGTGCAGTACCGCATGATGACCATCCTCTCCATGGCGACGG 3050
1561 GGGCACCCGGTGCAGTACCGCATGATGACCATCCTCTCCATGGCGACGG 1610
3051 CCTCAAGGTCCCGCTCTAGGGCCGTCTGATGTCATGGCGATTG.... 3096
1611 CCTCAAGGTCCCGCTCTAGGGCCGTCTGATGTCATGGCGATTGGATA 1660
3097 GGATATCATCCCGCTTAATCC..... TTAAAAAATT 3126
1661 TGGATATCGTCCCGCTTAATCCACGACAAATAACGCTCGTGTACAAATT 1710
3127 TGCATGCATGCATGTAAGGGAAAGCGATGGGTTTCATTGGTGGCTTGGCT 3176
1711 TGCATGCATGCATGTAAGGGAAAGCGATGGGTTTCATTGGTGGCTTGGCT 1760

10021657-132402

Figure 6D

3177 TAAGCCTTAAAAACTCCGTCGGGTCTGCGAACCAACCACATCACTAGTGT 3226
|||||||
1761 TAAGCCTTAAAAACTCCGTCGGGTCTGCGAACCAACCACATCACTAGTGT 1810
|||||||
3227 TTTGTACTCTACTCCTCAGTGGAAAGTGTAGTGACAGCATAACAAGTTCATC 3276
|||||||
1811 TTTGTACTCTACTCCTCAGTGGAAAGTGTAGTGACAGCATAACAAGTTCATC 1860
|||||||
3277 ATATATATTATCCTCTTCTTCGCCGGATGCTTCCGGGACCTTTGGAG 3326
|||||||
1861 ATATATATTATCCTCTTCTTAAAAAAAAAAAAAAACTCGAG.... 1906

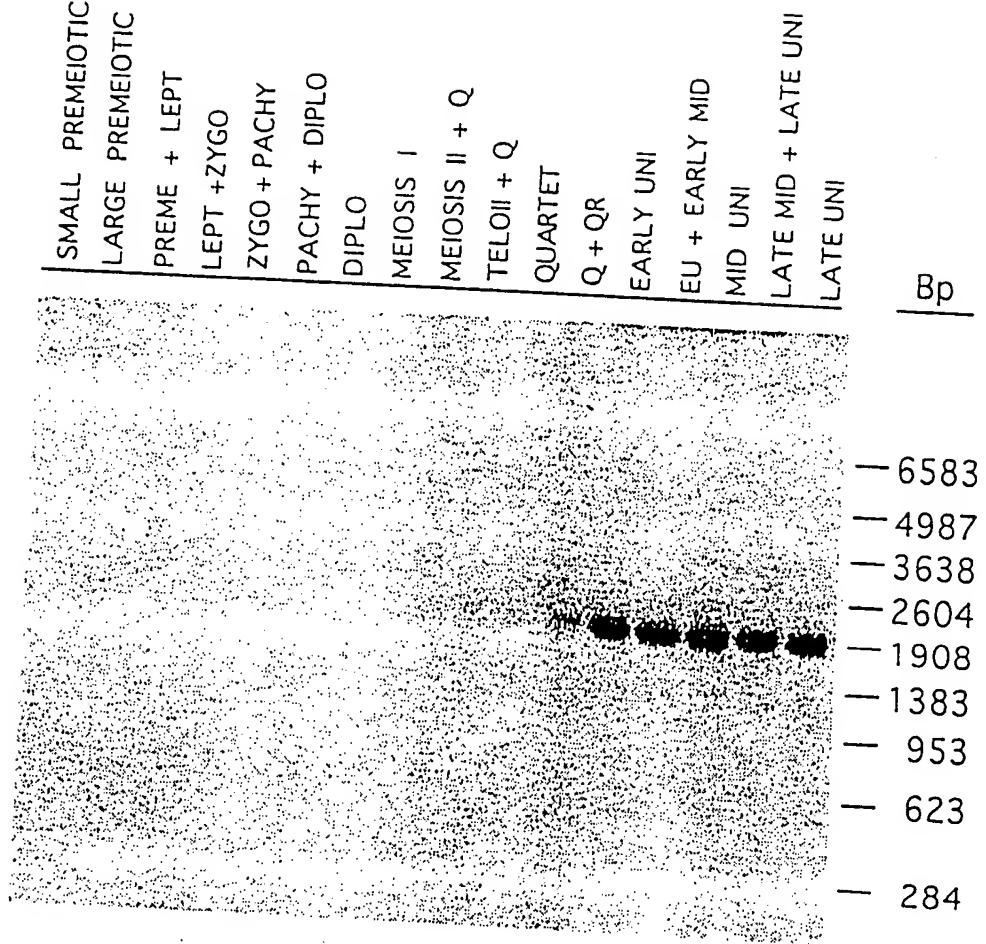


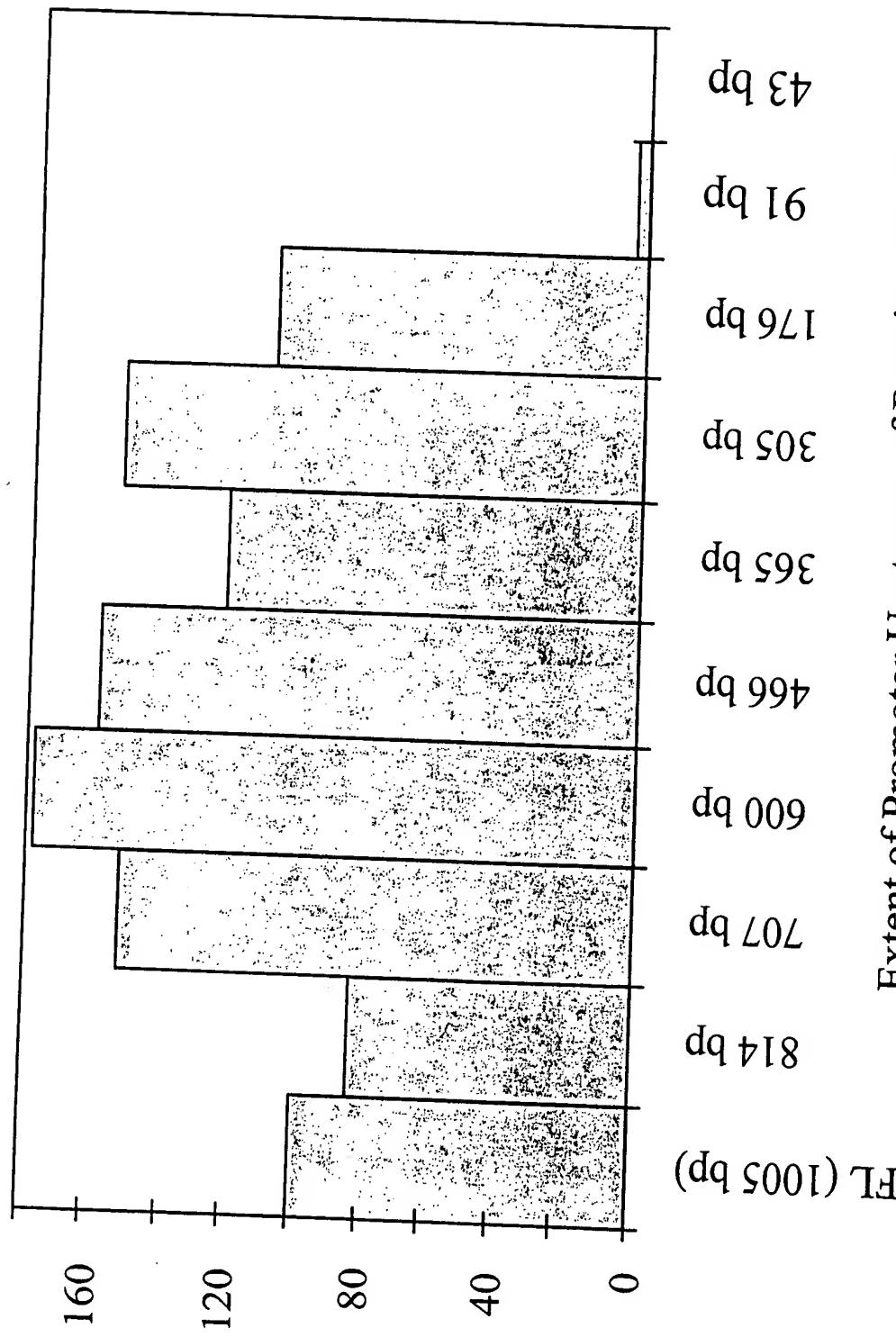
Figure 7

Figure 8

1 GAATTCCAAG CGAGGCCCTT GTAGCAGAGA GTGTTGCTGA TGCAGTCGGC
51 GGAAATGAGT GCGTGCTGAG AGCAACGCTG AGGGGTTCCA GGGATGGCAA
101 TGGCTATGGC AATCGGCTAG AGGTGGAGGA CAAGGTGGTG AGGATTGGGA
151 GGGCAACCTA TGGCAAGTTG GTGAAGAGGC ACGCAATGAG AGATCTATTC
201 AGACTTACAC TGGATGCCGC CAACAAATT AACCTTTAGA TTTTGATACT
251 GTCACTCCTA CTTTATTCCCT TGGTTGGCA ACTTCCAATA GGCTCATGTT
301 AATCAATGAT TAGTGATTAT TCAGCAAATA TTCTTGTGTTG TTTGACATTT
351 ATAATATGTG GGGTGAGACG GATTAAATAT CATCCATGAG AGCTTTATCT
401 TCATGCTCTC TTGATTTGG TTTCAGATCA TTCTTCAGT GTTCACAAGA
451 ATTTTCTCAG TTTGGTCCAT GTAATTGTTG AAGTGAGGTT CCTTAAATTT
501 CATTATGCTT CCTTTCTTTT CTAGACTAGC AACTGCATGA CTTTTCACCT
551 TGGGTTCACCA AATTGACTCA CAAGAAAACA AATTCACTTT TGGGTTCACCA
601 AATTCCCTTT CAGGATGTAC TTTCACTTG AACTGTCATG TATAGGAACA
651 AGGAATGGCT CAGTTTTAA GGAACAATGT ACAGATTCA TTTCAGAACT
701 CTTTCTGGTT GGTTGAGTTT CAGACTTTT GTACCAAGCT GATGGATCAC
751 AATACTTGTGTT TCCAAAGTCT GATAACAGAA ACTGGCAACT CCTAATTGAT
801 AATAAAAAGA ATAAAATACA GTATCAGATA TCTCATTTC TTGGTTGGCA
851 GATCACAAAA AGGAACACAA AGGCTAAGCC TCCTACTTGT TCGGGAGTTA
901 GGTCAGGGAC ACCATATGAA TGAAAGAAAT CTTAATTGG GGTACACACCA
951 AGATTGTCTC TCTCGAGGTT GGGGGTCCC TAAGGTTGGT AGTAGCAATA
1001 CCCAATATAT CACCTAACAA ACCCAATCCA TGCTACATAC ATACATAGCA
1051 TCCATCACTT GTAGACTGGA CCCTTCATCA AGAGCACCACGG

SBMu200 Promoter Analysis: 5' Deletions

Normalized Luciferase Activity as a
% of Full-Length Promoter Activity



Extent of Promoter Upstream of Putative TATA
Figure 9

SBMu200 “Minimal” Promoter

-180	CCCA <u>ATCTCA</u> TTTCTTGGT	TGGCAGATCA	CAAAAGGAA	CACAAAGGCT
	LS01	LS02	LS03	LS04
-130	<u>AAGCCTCC</u> TA CTTGTTCGGG	AGTTAGGTCA	GGGACACCAT	<u>ATGAAATGAA</u>
	LS06	LS07	LS08	LS10
-80	<u>GAATCTTA</u> TTTGGGGTCA	CACCAAGATT	GTCTCTCTCG	<u>AGGTTGGGG</u>
	LS11	LS12	LS13	LS14
-30	<u>GTCCCTAAGG</u> TTGGTAGTAG	<u>CAATACCAA</u>	TATAT <u>CACCT</u>	AACAAACCCA
	LS16	LS17	LS18	
20	ATCCATGCTA	CATACATACA	TAGCATCCAT	CACTTGTAGA
70	CATCAAGAGC	ACCATGG		CTGGACCCCTT

Coordinates are relative to the putative TATA box (underlined). P motifs are in italic.
□ = Del -176/-92 **□** = Del -89/-44 **□** = Del -39/-8
 Linker scanning mutations that reduce activity to ~5% or less are in bold. Mutations with a significant but less pronounced effect are in bold italic.

Figure 10

Linker Scanning Analysis of *SBMu200* Promoter

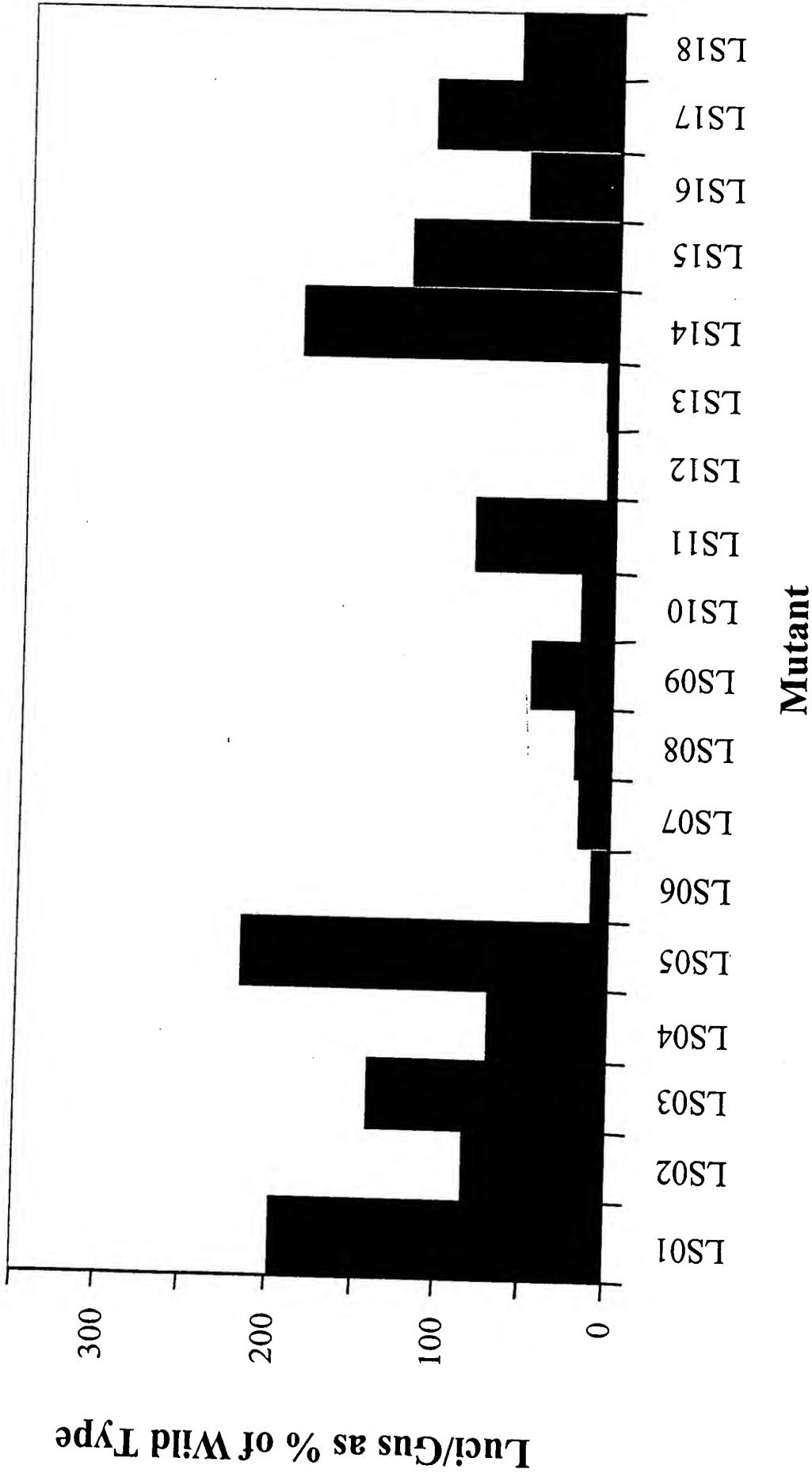


Figure 11

Figure 12

1 GGAA 4
201 CCGGATGCACGACTGGCTTGTGGTACCTGTACGGCACAGGACAGTGA 250
5 TTCGGCTTATGCCGTTCACTTCCTACACCTACATCGCTGACCCGGTGAAT 54
251 CCGTCGACATGCCGTTCACTTCCTACACCTACATCGCTGACCCGGTGAAT 300
55 GTCGAGCATGTCCCTCAAGACTAACTTCACCAATTACCCAAGGGGGACGT 104
301 GTCGAGCATGTCCCTCAAGACTAACTTCACCAATTACCCAAGGGAAATCGT 350
105 GTACAGATCCTACATGGATGTGCTCCTCGGTGACGGCATATTCAACGCTG 154
351 GTACAGATCCTACATGGACGTGCTCCTCGGTGACGGCATTTCAACGCCG 400
155 ACGGCGAGCTGTGGAGGAAGCAGAGGAAGACGGCGAGTTCGAGTTGCC 204
401 ACGGCGAGCTGTGGAGGAAGCAGAGGAAGACGGCGAGTTCGAGTTGCC 450
205 TCCAAGAACCTGAGGGATTCAGTGCCATGTTTCAGAGAGTACTCCCT 254
451 TCCAAGAACCTGAGGGATTCAGCGCCATTGTGTTCAAGAGAGTACTCCCT 500
255 GAAGCTGTGGGCATACTGAGTCAGGCATCCAAGGCAGGCAAAGTTGTTG 304
501 GAAGCTGTGGGTATACTGAGCCAGGCATCCAAGGCAGGCAAAGTTGTTG 550
305 ACATGCAGGAACCTTACATGAGGATGACACTGGACTCGATCTGCAANGTT 354
551 ACATGCAGGAACCTTACATGAGGATGACGCTGGACTCCATCTGCAAGGTT 600
355 GGGTCGGGTCNANATCGGCACGCTGTNCGGATCTCCCCGAGAACAG 404
601 GGGTCGGGTCGAGATCGGCACGCTGTCGCCAGATCTCCCCGAGAACAG 650
405 CTTCNCCAAGCGTTCGATGCCGCTAACATCATCGTCACNCTGCGGTTCA 454
651 CTTCGCGAGGCGTTCGATGCCGCAACATCATCATCACGCTGCGGTTCA 700
455 TCCACCCNCTGTGGCGATCCAGAAGTTCTTCCCCNGTCA..... 494
701 TCGACCCGCTGTGGCGATCAAGAGGTTCTTCCACGTCGGGTCAAGAGGCC 750

Percent Similarity: 92.510 Percent Identity: 90.891
Sb200-Sorghr.Pep x Sb20081.Pep February 13, 1997 11:29

5 MPFTSYTYIADPVNVEHVLKTNFTNYPKGDVYRSYMDVLLGDGIFNADGE 54
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
87 MPFTSYTYIADPVNVEHVLKTNFTNYPKGIVYRSYMDVLLGDGIFNADGE 136
||||| ||||| ||||| ||||| ||||| ||||| |||||
55 LWRKQRKTASFEFASKNLRDFSANVFREYSLKLSGILSQASKAGKVVDMQ 104
||||| ||||| ||||| ||||| ||||| ||||| |||||
137 LWRKQRKTASFEFASKNLRDFSIAIVFREYSLKLSGILSQASKAGKVVDMQ 186

Figure 12B

卷之三